Journal of the Lepidopterists' Society
39(3), 1985, 201–207

HOW TO DO GENETICS WITHOUT MAKING THE BUTTERFLIES CROSS

JOHN R. G. TURNER

Department of Genetics, University of Leeds,
Leeds LS2 9JT, England

ABSTRACT. It is possible to find out whether an inherited variety of a butterfly is sex-linked and, if it is not sex-linked, whether the variety is dominant or recessive to normal (in short, to find out its basic genetics) without carrying out pedigreed breeding experiments. These require much space, time and record-keeping, and are in any case not possible in some species.

Instead, one can raise offspring from the two types of female captured in the wild or followed while ovipositing. The mates of the females need not be observed. A fairly simple calculation based on the numbers of the two types of offspring produced by the two types of female will then reveal the inheritance of the variety.

The method is illustrated with data on the green and yellow forms of the African *Papilio phorcas*.

Working out how the different forms of a butterfly are inherited can be tedious; a pedigree record must be kept over a number of generations, the offspring of different females must be kept separate, and one needs to be fairly skilled in mendelism to set the crosses up in the way that will give the necessary information. If the variant one is studying is confined to the female, as are the white forms of some *Colias* or the black form of *Papilio glaucus,* the exercise becomes even harder, for as one cannot tell what color the male "ought" to be, one must do the crosses "blind." In addition, some butterflies cannot easily be mated in captivity.

It is, however, possible to do butterfly genetics without any of this hassle. Provided wild caught females can be persuaded to lay eggs or can be found ovipositing in the field, it is possible to determine the genetics of naturally occurring forms simply by raising the offspring of wild females. Neither the possibility that the female may be producing a mixed brood after mating with two males, nor even combining the offspring of different females in one breeding cage, will spoil the method. The only requirements are that one must be certain which color of female was the mother of the eggs, and that one of the forms should be rarer than the other. (When the forms are exactly equal in frequency the method fails completely and it requires very large numbers of offspring indeed when the rare form is over around 35% of the population.) With some tropical species one must be very cautious about information obtained from whole egg rafts, as these are sometimes laid cooperatively by several females; unless all the females have been seen from the laying of the first egg, and all are of the same color, the result can be completely unreliable (Turner, 1971, 1981; Mallett & Jackson, 1980).

I will describe the method for a butterfly having just two forms, and use data on the green and yellow forms of the African *Papilio phorcas* for illustration. North temperate zone workers may find it easier to think in terms of *Colias*: to do this, simply substitute mentally "white" for "green." With three or more forms the method becomes, needless to say, more complicated.

The method depends on a principle readily derived from the tenets of population genetics, that if the females of a rare form, having mated randomly with the males in their population, give rise to offspring which are mostly of the common form, then the rare form is recessive. On the other hand, if the rare form is dominant, it will give this fact away by producing among its offspring roughly equal numbers of the two forms. The common form, whether dominant or recessive, always tends to produce a majority of offspring like itself.

Some mathematical precision can be given to this idea (I give the proof elsewhere—Clarke et al., 1985). If the frequencies of the dominant and the recessive *genes* (not forms) in the population are p and q, then the recessive females, overall, produce offspring in just these proportions. If the green form of *Papilio phorcas* was recessive and the gene frequencies for yellow and green were 75% and 25%, then in aggregate a sample of eggs from a number of green females would produce 75% yellow and 25% green offspring. So, calling D the proportion of dominants in the offspring and R the proportion of recessives, we have the formula for the offspring of *recessive* females:

$$D = p, \qquad R = q \tag{1}$$

where p is the frequency of the dominant gene and q the frequency of the recessive.

Dominant females on the other hand produce the two forms according to the formula

$$\frac{D}{R} = \frac{p}{q} + \frac{1}{q^2} \tag{2}$$

where p and q are as before the frequencies of the dominant and recessive genes.

If the gene frequencies were as before, but the *green* form was dominant, then in the offspring of green females we would have

$$D/R = 0.25/0.75 + 1/0.75^2 = 2.11$$

and

TABLE 1. Offspring of wild females of *Papilio phorcas* from Nairobi and Ngong (Kenya). From Clarke et al. (1985).

Offspring	Mother Green	 Yellow
Green	78 (84%)	37 (51%)
Yellow	15 (16%)	35 (49%)

$$D = 2.11/(1 + 2.11) = 0.68, \qquad R = 1/(1 + 2.11) = 0.32$$

so that 68% of the offspring of green females would be green and 32% yellow.

To determine whether a rare form is dominant or recessive, we therefore compare the frequencies of the two forms among its offspring with those we would expect according to formula (1) and formula (2). Provided the butterflies are mating at random, one of these formulae will give an answer fitting the data, and the other will not. Obviously, to do this we need to know the values of p and q which, as we will see, can be obtained *either* from a population sample or from further breeding work.

What is needed, therefore, is a set of offspring derived from females of the *rare* form. There is no need for the mate of the mother to be known, nor to have any minimum number of offspring from any one female (they could well be eggs found by following ovipositing females around in the field), nor any need to keep the offspring of different females separate. All that is needed is the certainty that they are the offspring of the rare type of female.

In addition, it is necessary to have an estimate of the frequency with which the rare form occurs in the population, obtained by catching as many individuals as possible without making a special effort to capture either kind, or provided the population is large and the butterflies not too sedentary, simply by keeping a tally of the numbers of the two forms seen. If this is not obtainable, a satisfactory substitute is a large set of offspring derived from the *commoner* kind of female. Again, so long as they certainly are from this type of female, no further information is needed.

In sum, we need (1) a set of offspring from the *rarer* type of female, *plus* (2) *either* a field estimate of the proportions of the two forms *or* a set of offspring from the *commoner* type of female. Data of this kind for *P. phorcas* are shown in Table 1, where I have combined all the offspring of a large number of wild green mothers and of yellow wild mothers from the Nairobi area (including the town of Ngong). In addition, the yellow form has been reported as rare in this region, probably being a little less than 20% of the population. Suppose first that yellow is recessive. The frequency of the gene is then given by

$$q = \sqrt{0.2} = 0.447$$

(a surprisingly large frequency—nearly 45%—as recessive genes are always much more common than the form which they control). As the yellow form is recessive, yellow females should produce yellow and green offspring, from formula (1), in the proportions

$$\text{yellow} = q = 0.447$$
$$\text{green} = p = 1 - q = 0.553$$

These proportions are close to the observed numbers of the two kinds of offspring and we strongly suspect that the yellow form is recessive.

Does the hypothesis that the yellow form is dominant fare worse? In that case the frequency of the green gene (which must be recessive) is

$$q = \sqrt{(1 - 0.2)} = 0.894$$

As the yellow form is dominant, it will give rise to yellow and green forms, according to formula (2), in the ratio

$$\text{yellow}/\text{green} = p/q + 1/q^2$$

or in this case 1.368 : 1. This means that among the offspring we expect

$$1.368/(1 + 1.368) = 0.578 \text{ yellow and}$$
$$1/(1 + 1.368) = 0.422 \text{ green}$$

which is not such a good fit to what is actually observed (Table 1). The yellow form therefore appears to be recessive.

However, suppose that we do not have a good estimate of the frequencies of the two forms in the population (and the estimate of 20% yellow is in fact not particularly accurate). A perfectly good substitute for this estimate is the number of the two forms appearing among the offspring of the *common* female form. Our data for the numbers of yellow and green females arising from green mothers are given also in Table 1.

Start by supposing that yellow is dominant. In that case the yellow and green proportions from the *green* mothers are direct estimates of the gene frequencies p and q, giving in this case $p = 0.161$ (yellow) and 0.839 (green). We can test this against the offspring of yellow females, again by using the formula

$$\text{yellow}/\text{green} = p/q + 1/q^2$$

and in this case yellow: green is 1.614:1; yellow individuals should be $1.614/(1 + 1.614) = 0.617$ and green individuals $1/(1 + 1.614) = 0.383$ of the offspring. Again, the fit is not very good.

Checking whether yellow being recessive gives a good fit is harder this time. If yellow is recessive the ratio of green/yellow from green mothers, which from Table 1 can be calculated as $78/15 = 5.2$, will give q if we solve the equation

$$(1 - q)/q + 1/q^2 = 5.2$$

This is a quadratic in q, and according to standard algebra, the general solution is that if x is the ratio of green to yellow from green mothers, then

$$q = \frac{1 \pm \sqrt{4x + 5}}{2x + 2} \tag{3}$$

Substituting 5.2 for x in (3) gives us $q = 0.490$ and therefore $p = 0.510$. These should be the proportions of yellow and green among the offspring of yellow mothers, which is clearly an excellent fit (Table 1). The yellow form is obviously recessive.

The results, particularly if numbers are small, might not be so obvious as this, and then a statistical test would have to be applied, comparing the observed and expected *numbers* (*not* the percentages).

We can summarize the value of the method with Table 2. The first column shows the frequency of the *form* which is actually recessive, and the next the frequency of the recessive gene. If we obtained offspring from recessive females (which are the rare form *above* the line and the commoner form *below* it) we would obtain the offspring proportions shown in the third column; the fourth column shows the offspring which would be obtained from the dominant females (which are the rare form in the lower half of the table). The last column shows the proportions which we would calculate for the offspring of rare females (recessive above the line, dominant below) when we took the *wrong* hypothesis about the dominance. By comparing this with the numbers in bold type, we can see how easy, or not, it is to tell that we are in fact wrong. It can be seen that provided one or other form is below about 30%, the method will distinguish very well which of the forms is dominant but that it will not work when the forms are nearly equally common in the population.

It is, however, still worth making the observations even when the forms are equally abundant, for this allows us to distinguish a sex-linked gene. For if the gene were carried on the X chromosome, then *both* kinds of female would produce offspring in the same proportions:

TABLE 2. Proportions of two types of offspring from both types of mother at different population frequencies of the rare form.

Actual frequency of recessive form (yellow)	Actual frequency of recessive gene (yellow)	Recessive (yellow) females give dominant : recessive (green : yellow)	Dominant (green) females give dominant : recessive (green : yellow)	Wrong assumption about dominance of rare form predicts that it will give green : yellow
0.001	0.032	**0.968:0.032**	0.999:0.001	**0.499:0.501**
0.005	0.071	**0.929:0.071**	0.995:0.005	**0.496:0.504**
0.01	0.100	**0.900:0.100**	0.991:0.009	**0.493:0.507**
0.05	0.224	**0.776:0.224**	0.959:0.041	**0.470:0.530**
0.1	0.316	**0.684:0.316**	0.924:0.076	**0.444:0.556**
0.2	0.447	**0.553:0.447**	0.862:0.138	**0.399:0.600**
0.3	0.548	**0.452:0.548**	0.806:0.194	**0.360:0.640**
0.4	0.633	**0.367:0.633**	0.755:0.245	**0.325:0.675**
0.5	0.707	**0.293:0.707**	0.707:0.293	**0.293:0.707**
0.6	0.775	0.225:0.775	**0.662:0.338**	**0.633:0.367**
0.7	0.837	0.163:0.837	**0.619:0.381**	**0.548:0.452**
0.8	0.894	0.106:0.894	**0.578:0.422**	**0.447:0.553**
0.9	0.949	0.051:0.949	**0.538:0.462**	**0.316:0.684**
0.95	0.975	0.025:0.975	**0.519:0.481**	**0.224:0.776**
0.99	0.995	0.005:0.995	**0.504:0.496**	**0.100:0.900**
0.995	0.998	0.002:0.998	**0.502:0.408**	**0.071:0.929**
0.999	0.9995	0.0005:0.9995	**0.5004:0.4996**	**0.032:0.968**

The ease with which one can tell which form is in fact recessive can be seen by comparing, in any particular row, the figures printed in bold type. Within the dotted lines, the figures are well-matched and the dominance is hard to determine; above and below these lines there is clear discrimination, and this is particularly marked when the recessive form is very rare or very common, as at the top and bottom of the Table.

say 60:40 green and yellow from both green and yellow mothers. Whereas, if the gene is not on the sex chromosome, the proportions, as can be seen from the center line of Table 2, are mirror images; the yellow form produces yellow : green in the ratio 0.71:0.29, whereas green produces them in the ratio 0.29:0.71.

It should be noted that this method becomes *completely* unreliable if the offspring of pedigreed captive matings are included in the data; the only permissible use of captive bred butterflies is to take virgin females and mate them to wild-caught males, or to collect larvae at random in the wild and then test their offspring, for the first generation only, by mating them in captivity. Indeed, when I first tried to apply the method to *Papilio phorcas,* there were few matings and I included the offspring of some pedigree broods to swell the numbers; the method then gave the totally incorrect answer that yellow was dominant, which shows how unreliable it is in those circumstances.

I believe that useful information could be obtained on the genetics of some of the more "difficult" species of butterflies and moths, by using this technique. The recipe provided above will be found quite easy to follow if it is applied step by step. As an example, readers might like to try to determine the inheritance of an imaginary white *Colias,* occupying 3% of its population, and giving 51 white to 49 orange from white mothers.

As a matter of history, it is worth recording that the first use of a primitive version of this method appears to have been by E. B. Poulton (1914), who determined in this way that one of the rare forms of *Papilio dardanus* was produced by a dominant gene.

ACKNOWLEDGMENTS

I am most grateful to Professor Sir Cyril Clarke, KBE, FRS and to Lady Clarke, who suggested this problem and the broad outline of its solution to me and who provided the data on *Papilio phorcas.* They also read the draft and suggested some improvements.

LITERATURE CITED

CLARKE, C. A., F. M. M. CLARKE, S. C. COLLINS, A. C. L. GILL & J. R. G. TURNER. 1985. Male-like females, mimicry and transvestism in butterflies (Lepidoptera: Papilionidae). System. Entomol. 10:257–283.

MALLET, J. L. B. & D. A. JACKSON. 1980. The ecology and social behaviour of the neotropical butterfly *Heliconius xanthocles* Bates in Colombia. Zool. J. Linn. Soc. London 70:1–13.

POULTON, E. B. 1914. The Mendelian relationship of the female forms of *P. dardanus.* Proc. Entomol. Soc. London 1914, lxvii–lxx.

TURNER, J. R. G. 1971. Studies of müllerian mimicry and its evolution in burnet moths and heliconid butterflies. *In* E. R. Creed (ed.) Ecological genetics and evolution, pp. 224–260. Oxford, Blackwell.

——— 1981. Evolution and adaptation in *Heliconius*: a defence of neo-Darwinism. Ann. Rev. Ecol. Syst. 12:99–121.